

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

BNSDOCID: <WO 9916904A1 | >

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

EFFICIENT LINKING OF NUCLEIC ACID SEGMENTS

BACKGROUND OF THE INVENTION

Technical Field

This invention relates to a method for the efficient amplification of target DNA segments which is particularly advantageous for those target genes containing many small exons. Specifically, the method involves a rapid polymerase chain reaction technique for linking multiple gene segments from a single gene or multiple genes into a large DNA molecule suitable for further analysis.

Description of the Background Art

The technique known as the polymerase chain reaction (PCR) is a method of amplification of genomic DNA (Saiki et al., Science: 1350-1354 (1985)). Typically, the method uses two oligonucleotide primers to amplify a single DNA segment millions of times. Each cycle of DNA replication from the original primer(s) produces a product which serves as a template for further primer-dependent replication. This feature of the method results in exponential increases in the desired DNA product with each round of amplification, and a rapid accumulation of DNA. Now an automated technique performed using a thermal cycler and

thermostable DNA polymerases, PCR is widely used by molecular biologists to prepare large amounts of synthetic DNA.

5 This revolutionary technique has been used in a wide variety of fields in molecular biology, and has made possible the rapid identification of disease-associated genes. Using the PCR, it has become feasible to diagnose inherited disorders and susceptibility to disease at the molecular level.
10 Nevertheless, several disadvantages and limitations are recognized in the technique and its application to certain genes.

Customarily, when using PCR to amplify genomic DNA, each gene segment is amplified separately and then
15 analyzed. When the gene of interest contains more than one exon, each exon must be amplified individually using a separate PCR, and then linked together to form a long DNA molecule representing the entire gene (Ho, et al., 1989; Horton, et al., 1989). Because no more
20 than two gene segments can be linked together in each joining PCR, the more exons the gene of interest contains, the more separate amplifications must be individually performed and the more PCRs are needed to link the segments together. For genes which contain a
25 lot of small exons, the preparation of a single gene for analysis can become prohibitively labor-intensive, and require a great deal of time, particularly when the small target exons must be individually scanned for mutations or polymorphisms.

30 It has been shown that simultaneous amplification of more than one DNA segment can be achieved with a Multiplex PCR using primers tagged with an unrelated 20

nucleotide sequence from bacteriophage M13mp18 (Shuber
et al., 1995). However, with this method, products
amplified with primers lacking the 20 nucleotide
sequence were not reliably produced due to differences
in hybridization kinetics among the primers. Using the
prior art method, therefore, tagging each primer with
an identical 20 nucleotide sequence is necessary to
achieve efficient amplification of multiple sequences.
This prior art method thus allows multiple
amplifications, but the products of the amplification
all contain identical unrelated sequences which would
have to be removed or extended before they could be
linked to form one, long DNA molecule containing all
portions of the gene of interest.

After the individual gene segments have been
separately amplified, further independent steps are
needed to reconstruct the complete desired gene
sequence from the smaller segments, with or without an
introduced mutation, before the entire gene is ready
for analysis. Prior art methods for linking sections
of DNA using PCR involve the joining of two segments at
a time, each PCR followed by a purification step. (Ho
et al., 1989; Kim et al., 1996). The joining of
several gene segments together, therefore, requires
multiple PCRs and multiple purification steps. For
example, joining four exons to form one complete gene
using these prior art methods would require four
separate amplifying PCRs, four separate purifications
of the products and three joining PCRs. Each
additional DNA segment in the gene would require an
additional PCR to link it to the others, increasing

both the time and expense of preparing the DNA for analysis using prior art methods.

In summary, the prior art methods of producing an amplified copy of an entire gene composed of multiple linked exons or an amplified copy of a long DNA molecule composed of gene segments from more than one gene have the disadvantage that multiple PCRs are generally required at each step. Previously, it has not been possible to efficiently amplify multiple gene segments in a single PCR to yield products that had complementary ends suitable for easy, rapid linkage in a second single PCR.

Consequently, there has been a need in the field for a simple and rapid method allowing amplification of DNA which contains several linked DNA segments which occur in non-adjacent portions of target DNA. There is a need for a method which can produce an amplified DNA molecule containing an entire gene of linked exons from genomic DNA, e.g., for DNA diagnosis.

SUMMARY OF THE INVENTION

The present invention provides a method of linking by PCR DNA segments which occur in non-adjacent portions of target DNA wherein each DNA segment contains a sequence complementary to a sequence in the DNA segment or segments to which it is to be linked, comprising using a) a first primer which is complementary to the antisense strand of the first DNA segment to be linked and a second primer which is complementary to the sense strand of the last DNA segment to be linked; and b) at least one polymerase

lacking 3'-5' exonuclease activity and at least one polymerase containing 3'-5' exonuclease activity.

In addition, the present invention provides a method of producing and amplifying DNA containing at least three linked DNA segments which occur in non-adjacent portions of target DNA, comprising a) providing a first primer and a second primer for each DNA segment to be amplified, i) the first primer (termed the D primer) having a 3' portion which is complementary to the 3' end of the antisense strand of the DNA segment and a 5' tail which is complementary to the 5' end of the second primer for the previous DNA segment or to a sequence internal to the previous DNA segment; ii) the second primer (termed the U primer) having a 3' portion which is complementary to the 3' end of the sense strand of the DNA segment and a 5' tail which is complementary to the 5' end of the first primer for the subsequent segment or to a sequence internal to the subsequent DNA segment; b) amplifying the at least three DNA segments by multiplex PCR using the pairs of first and second primers; and c) subjecting the at least three amplified DNA segments to a linking PCR using a sense primer which is complementary to the antisense strand of the first segment to be linked and an antisense primer which is complementary to the sense strand of the last segment to be linked.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A is a schematic representation of the Multiplex and Linking PCR steps of the inventive process. Four regions of the p53 gene were amplified by Multiplex PCR with the four primer pairs D₁/U₁, D₂/U₂, D₃/U₃ and D₄/U₄. Each primer contains a GC-rich tail and a sequence-specific region. The tail of a U primer is complementary to the tail of subsequent D primer. After a simple purification step, the four PCR amplified DNAs (D₁U₁, D₂U₂, D₃U₃ and D₄U₄) are linked and amplified by nested P and Q primers.

Figure 1B provides an example showing two types of tails which can be used with the inventive process. The type I tail of D₃ primers is not overlapped with the sequence-specific region of U₂ primer, while the Type II tail is overlapped by 4 bases. Sense (SEQ ID NO:21) and antisense (SEQ ID NO:22) sequences of the p53 gene, and sense (SEQ ID NO:23) and antisense (SEQ ID NO:24) sequences of the F9 gene are included in the tail.

Figure 1C. This schematic diagram illustrates the use of the type III tail, in which the tail sequence is complementary to an internal portion of the DNA sequence to be joined.

Figure 2A shows the relative yields of PCR product of the p53 gene with varied amounts of *Vent* and fixed amounts of *Tth*, *Taq*, or *Tfl* enzymes. Fixed amounts of *Tth*, *Taq*, or *Tfl* and increasing amounts of *Vent* were used to link and amplify segments of the p53 gene. Relative yields of the linked PQ product were quantitated using a PhosphorImager (Molecular Dynamics) after 15 cycles.

Figure 2B shows the relative yields of PCR product of the F9 gene with varied ratios of *Tth* and *Vent* enzymes. Increasing amounts of *Tth* and *Vent* were used to link and amplify segments of the F9 gene. Again, relative yields of the linked PQ product were quantitated using a PhosphorImager after 15 cycles (M=120ng ϕ x174/Hae III DNA marker).

Figure 3 demonstrates the effect of differing amounts of DNA template on the yield of PCR product. In Figure 3A, the p53 gene was amplified using *Tth/Vent* DNA polymerases. *Tth/Vent* DNA polymerases were applied with increasing amounts of template DNA and the p53 PQ product quantitated. In Figure 3B, *Tth/Vent*, *Taq/Pfu* or *Tfl/Vent* DNA polymerases were applied to the F9 gene. *Tth/vent*, *Taq/Pfu* or *Tfl/Vent* were applied to the F9 gene.

Figure 4 presents the relative yields and accumulation of PCR product. Aliquots of identical radioactively labeled Linking PCR mixtures were removed from the thermocycler every 3 cycles from 9 to 30 cycles and the PQ PCR product was quantitated. Forty nanograms (4A), 20ng (4B), or 10ng (4C) of p53 DNA templates per 25 μ l reaction were used.

Figure 5 gives the relative yields of PCR products of the p53 and F9 genes with different annealing temperature.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The gene amplification method of the present invention can produce large amounts of DNA composed of several exons or a DNA composed of several non-

contiguous DNA segments from the same gene or different genes, without requiring the time-consuming amplification of each separate gene segment. Another advantage of the present invention is the easy, one-
5 step linkage of the gene segments and amplification of the linked product.

The invention has multiple uses which include but are not limited to, the following: i) efficient scanning of mutations by methods such as restriction
10 endonuclease fingerprinting when genomic DNA is analyzed from genes in which there are multiple short exons separated by long introns; ii) joining of different protein domains to generate a recombinant gene/RNA which has novel properties; and iii) linking
15 RNAs together by generating cDNA, linking the cDNA with the primer that contains an RNA promoter sequence, and after linkage transcribing the linked segment to generate the RNA.

20 Multiplex PCR

Multiplex PCR is the amplification of the desired regions (for example, exons) of the genetic material using a pair of primers for each individual region. Figure 1 illustrates in schematic form the
25 amplification of four exons simultaneously with four primer pairs prior to the linking step. The primer pairs are designated D_1/U_1 , D_2/U_2 ... D_n/U_n . Each primer contains a GC-rich tail and a sequence-specific region, and the tail of each U primer is complementary to the
30 tail of the subsequent D primer. Three types of primer tails are contemplated for use with the invention. In type I primers the tail of the D primer does not

overlap with the sequence-specific region of the previous U primer. In type II primers, the tail of the D primer overlaps the sequence-specific region adjacent to the previous U primer. See Figure 1B. Type III primers contain a tail portion which is complementary to a sequence internal to the previous gene segment. See Figure 1C for an example in schematic form.

Primers were designed with Oligo 5 software (National Biosciences, Inc.) and the GCG program (Genetic Computer Group, Inc.). Oligo 5 calculates primers' melting temperature (T_m) by the nearest neighbor method at 50 mM KCl and 250 pM DNA. the T_m value of each PCR DNA was estimated by the Wetmur formula ($T_m^{\text{product}} = 81.5 + 16.6 \log[K^+] + 0.41(\%G + \%C) - 675/\text{length}$ (Wetmur, 1991). Type I tails do not overlap the sequence-specific region of the complementary primer, while type II tails overlap the sequence-specific region for four bases. Tails which overlap by more or fewer bases are also suitable for use with the invention. Type III tails are complementary to an internal sequence within the previous gene segment.

Guidelines for Primer Design

Designing the appropriate primers is a critical step in successfully performing Linking PCR. Based on this work, building on other studies using Multiplex PCR (Liu et al. 1997), the following guidelines for primer design were developed and successfully applied.

a. Sequence-specific region

The sequence-specific region affects the yields and specificities of the Multiplex PCR. The criteria are set as follows:

1. The T_m value should be approximately 35°C below the average T_m value of the targeted regions. A T_m value lower than this may result in low PCR yields, especially if the region of the gene segment to be amplified contains a high GC percentage.

2. The stringency for dimer or hairpin formation at the 3' end should preferably be set at ≤ 4 base pairs among all primers. This has the potential to cause a greater problem in Multiplex PCR than in ordinary PCRs using only two primers.

3. The stringency for false priming sites at the 3' end should preferably be set at ≤ 6 base pairs for all strands and for all regions.

4. Internal stability may be chosen based on the instructions in the Oligo 5 software package.

b. Tail region

The tail is short (preferably less than 20 bases) and contains a high percentage of GC bases, which functions to provide consistent and balanced high yields of Multiplex PCR products, and an efficient and specific "linker" for the Linking PCR. The criteria should be set as follows:

1. The GC content should preferably be from 60% to 70%.

2. The tail size is preferably 10-15 bases long and most preferably 12 bases long.

3. The stringency for false priming of the primer's antisense sequence at its 3' end should preferably be ≤ 6 bases for any strand and any target.

4. A type II tail is preferred.

Optimization of Multiplex PCR

The parameters of the PCR may be optimized according to Shuber, et al. (Shuber et al., 1995) or determined empirically. For the following examples, the optimization strategy of Shuber, et al. (Shuber et al., 1995) was followed, except for the annealing temperature. The optimal annealing temperature was determined empirically and was expected to be approximately 20-25°C below the average T_m of the gene regions being amplified. (Liu et al., 1997). A preferred strategy for optimization of the Multiplex PCR step is as follows:

a. Test each PCR:

Concentrations of primer, Mg, DMSO, and the amount of TaqGold DNA polymerase should be optimized for each polymerase chain reaction. The optimal annealing temperature should be approximately 20-25°C lower than the average T_m of the regions to be amplified, but ultimately should be determined empirically. If a region is not being efficiently amplified, adding an additional one or two bases to the sequence-specific region of the primer may increase the yield.

b. Test the Multiplex PCR:

The common parameters of each PCR should be chosen to generate balanced high yields of the specific desired Multiplex PCR products. The Taq DNA polymerase may be present in amounts as high as 2-6 units per 25 μ l reaction. Rarely, the primer concentration may need further adjustment to achieve even, balanced yields of each DNA segment. If satisfactory results are still not achieved, a change in the primer sequence may be necessary.

5 TaqGold with hot-start was found important to
prevent primer dimer formation and false priming. One
potential difficulty in the *p53* gene was the
amplifications of exons 10 and 11, which are separated
by an intron of 800 base pairs. However our results
showed no large PCR DNA spanning the two exons when the
10 T_m of the sequence-specific regions of the U_3 and D_4
primers was increased and their relative concentrations
adjusted according to our preferred optimization
scheme.

Linking PCR

15 Amplified DNA segments produced by multiplex PCR
or any other suitable method may be joined with the
linking PCR method, with or without prior purification
to remove unincorporated primers. First, the anti-
sense strands of U_1 and D_2 tails, U_2 and D_3 tails, and U_3
and D_4 tails are annealed and extended, so the four
 D_1U_1 , D_2U_2 , D_3U_3 , and D_4U_4 DNAs are linked into a D_1U_4
20 molecule in numerical order. If the primers are
complementary to a different region of the DNA segment
to be joined, the complementary regions are annealed
and extended. Second, PCR amplifies the joined
template with nested primers such as P and Q. (Figure
25 1A).

30 Tails of 12-base size worked efficiently, although
tails of 10-15 bases, or a greater range, are also
suitable. The tails of primers P and Q prevent
"megapriming," which occurs when a PQ product generated
in an earlier cycle acts as a primer for a larger D_1U_4
template in a subsequent cycle (Sarkar and Sommer,
1992; Sarkar and Sommer, 1990). Also, the tail acts as

a switch from low amplification efficiency to high efficiency, depending on which template of D₁U₄ or PQ to which the primer anneals (Liu, et al. 1997).

One of three DNA polymerases lacking 3'-5' exonuclease activity (*Tth*, *Taq* (Boehringer Mannheim) and *Tfl* (Promega)) were combined with one of two enzymes possessing 3'-5' exonuclease activity (*Vent* (New England BioLabs), *Pfu* (Stratagene)) to perform the inventive method. The effect of the enzyme which lacks 3'-5' exonuclease activity is speculated to remove the potential extra non-template A base at the 3' end of the PCR product (Wu et al., 1989). Persons of skill in the art will recognize that other enzymes may be used with the present invention, such as *Pwo* and *Plo*, but it is key that the polymerase activity is due to one (or more) enzymes without 3'-5' exonuclease activity and one (or more) enzymes with 3'-5' exonuclease activity. These enzymes and enzyme combinations serve only as examples by way of illustration and are not intended to limit the invention.

A solid or liquid macromolecular additive may be used in the linking PCR mixture. Macromolecular additives such as polyethylene glycol (PEG) may reduce the amount of template needed to obtain a satisfactory result.

Optimization of Linking PCR

The following preferred parameters are not intended to limit the invention. Skilled molecular biologists will recognize that different parameters may be used with the invention.

a. Primers

The nested primers P and Q should be designed using the same criteria and methods as described above for the D and U primers, except that the T_m of the sequence-specific regions are preferably approximately 35-40°C lower than the D_1U_4 DNA product.

b. Polymerases

Tth/*Vent* DNA polymerases are preferably present at approximately 1U/0.1U or 1U/0.05U per 25 μ l reaction.

c. Linkage efficiency

The linkage of individual DNA segments is preferably tested by measuring the linking efficiencies of all the regions desired to be linked, and all shorter linked segments. For example, if the desired complete DNA sequence is made up of 4 segments, the linkage efficiency of 4, 3, and 2 segments would be measured with the appropriate primer pairs. Table 1 illustrates this suggested method.

Table 1. Relative mole ratios of PCR products with different primer pairs

Gene	Template amount	Primer pair							
		D_1/U_2	D_2/U_3	D_3/U_4	D_1/U_3	D_2/U_4	D_1/U_4	P/Q	None
p53	40ng	10.85	8.10	9.09	6.73	7.27	4.81	5.31	0.22
	20ng	7.74	4.56	7.60	0.98	3.19	0.57	1.00	0.17
	10ng	8.54	4.44	7.08	1.07	0.56	0.09	0.13	0
F9	40ng	18.54	18.58	17.60	14.43	14.27	8.77	4.87	0.16
	10ng	17.78	17.60	16.85	8.18	8.58	3.61	1.00	0

The mole ratio of D_1U_2 , D_2U_3 , D_3U_4 , D_1U_3 , D_2U_4 , and no primer to D_1U_4 primer is obtained by normalizing the relative yield by the potential amount of incorporated radioactive 32 P-dCTP.

d. Annealing temperature

The optimal annealing temperature should be determined using a large amount of DNA templates, and is generally associated with the percentage of GC bases in the DNA templates. See Figure 5.

e. DNA template concentration

The preferred DNA template concentration is determined as shown in Figure 3, and is dependent on how many DNA segments are to be linked together.

f. Cycle number

The optimal cycle number for each Linking PCR should be determined. Routinely, 20-25 cycles are most efficient and yield the best product.

g. Quality control

The identity and quality of the Linking PCR product is preferably confirmed by direct sequencing. If the product is not of the correct sequence, the tails of the primers from the multiplex step and the P and Q primers should be double-checked.

EXAMPLES

1. Multiplex PCR of p53 gene exons

Each of four primer pairs of D_1/U_1 , D_2/U_2 , D_3/U_3 , and D_4/U_4 (Table 1A) were used to amplify exons 1, 2-4, 10 and 11 in the p53 gene. Each primer contained a GC-rich tail and a sequence-specific region. The tails of the U primers were complementary to the tails of each subsequent D primer. This example used a type I tail, in which the tail of the D_3 primer is not overlapped with the sequence-specific region of the U_2 primer. (Table 1A, Figure 1A). A hot-start at 92°C for 10

minutes was included for enzyme activation.. The denaturation was at 94°C for 15 seconds, and the annealing was at 55°C for 30 seconds, followed by elongation at 72°C for 2 minutes, for a total of 35 cycles with a Perkin Elmer model 9600 thermal cycler. The PCR mixture contained 50 mM KCl, 10 mM Tris/HCl, pH 8.3, 1.5 mM MgCl₂, 200 µM of each dNTP, 5% DMSO, 3-4U of TaqGold DNA polymerase (Perkin Elmer), and 250ng of genomic DNA per 25µl of reaction. After purification in a Centricon^R-100 microconcentrator (Amicon), the amount of DNA was determined by spectrophotometer at 260 nm. The four expected DNA products were obtained in similar molecular ratio, and the complementary tails did not cause obvious problems.

2. Multiplex PCR of F9 gene exons

Each of four primer pairs of D₁/U₁, D₂/U₂, D₃/U₃, and D₄/U₄ (Table 1B) were used to amplify exons 1, 2-3, 4, and 5 of the F9 gene. As in the above example, each primer contained a GC-rich tail and a sequence-specific region, and the tails of the U primers were complementary to the tails of each subsequent D primer. This example uses a type II tail, in which the tail of the D₁, D₂, and D₃ primers overlapped for four bases the sequence-specific region of the corresponding U primers. (Table 1B, Figure 1B). The PCR mixture and reaction parameters were the same as in example 1 with the exception that 5% DMSO was omitted from the reaction mixture.

Table 1. Primers

A.

	#	Name ^a	Conc (μ M)	Sequence ^b	tT_m^c ($^{\circ}$ C)	cT_m^c ($^{\circ}$ C)
p53	1	(5'UT) (750) 28D ₁	0.05	<u>gttcgcagagcgt</u> TTGTGCCAGGAGCCT	28.9	46.1
	2	(I1) (995) 29U ₁	0.05	<u>aggacgaccgct</u> AGCCCGTGACTCAGAGA	31.8	44.9
	3	(I1) (11641) 30D ₁	0.1	<u>agcggctcctcct</u> CTAGGGTTGGAAGTGTCT	31.8	43.3
	4 ^d	(I4) (12352) 30U ₁	0.1	<u>cgatggcacagc</u> GATACGGCCAGGCATTGA	30.9	51.5
	5	(I9) (17480) 31D ₁	0.06	<u>gctgtgccatcg</u> TCCGTCATAAAGTCAAACA	30.9	43.2
	6	(I10) (17741) 31U ₁	0.2	<u>gaggtgggtgtc</u> CCTATGGCTTTCCAACCTA	21.6	46.9
	7	(I10) (18547) 31D ₁	0.2	<u>gacacccacctc</u> ACCCTCTCACTCATGTGAT	21.6	42.7
	8	(I11) (18756) 29U ₁	0.06	<u>cccgtaggaca</u> GACCCAAAACCCAAAAT	29.2	43.9
	9	(5'UT) (769) 29D[P]	0.05	<u>caacgggtcagg</u> AGGGGTTGATGGGATT	29.5	43.3
	10	(I11) (18717) 29U[Q]	0.05	<u>cgtgtgggtgct</u> GAGGGAGGCTGTCAGTG	29.6	44.6
	#	Name ^a	Conc (μ M)	Sequence ^b	tT_m^c ($^{\circ}$ C)	cT_m^c ($^{\circ}$ C)
p53	11	(5'UT) (-102) 25D ₁	0.1	<u>tccgagag</u> GAGGCCATTGGAAATA	-7.2	41.9
	12	(I1) (248) 25U ₁	0.1	<u>gtaagcgc</u> CGTGCTGGCTGTTAGA	30.7	41.5
	13	(I1) (6104) 29D ₁	0.2	<u>cacgaccgc</u> TTACTGGAATTCTCTTGAAT	30.7	40.2
	14	(I3) (6870) 27U ₁	0.2	<u>gcccaacag</u> TGGCATAACCCTGTAGTAT	29.6	41.5
	15	(I4) (10264) 25D ₁	0.14	<u>gccactgt</u> TGGCTTCCAGGTCAGTA	29.6	43.1
	16	(I4) (10618) 28U ₁	0.14	<u>ggtcccgag</u> ATCAAAGGTATGTTTTTAAG	32.3	38.5
	17	(I4) (17584) 27D ₁	0.1	<u>tgatcccg</u> GACCCATACATGAGTCAGT	32.3	40.4
	18	(I5) (17897) 27U ₁	0.1	<u>acggagac</u> AGGAAGCAGATTCAAGTAG	-12.0	40.6
	19	(5'UT) (-57) 23D[P]	0.05	<u>tcaagag</u> GAGGGAGATGGACAT	-17.4	33.6
	20	(I5) (17856) 23U[Q]	0.05	<u>tggtgtgc</u> TAAAATGCTGAAGT	-17.1	27.2

^ap53 gene PCR products are D₁U₁ (270 bp, 58% G+C), D₂U₁ (736 bp, 59% G+C), D₃U₁ (286 bp, 55% G+C), D₄U₁ (235 bp, 54% G+C) and PQ (1433 bp, 57% G+C). The numbering system is based on GenBank Accession: X54156. F9 gene PCR products are D₁U₁ (367 bp, 40.6% G+C), D₂U₁ (784 bp, 31.4% G+C), D₃U₁ (371 bp, 39.9% G+C), D₄U₁ (330 bp, 36.4% G+C) and PQ (1702 bp, 35% G+C). The numbering system is as described in Yoshitake, et al. (Yoshitake, et al., 1985). Key to primer names: 5'UT indicates the primer begins at a 5'

5 untranslated region; the letter I followed by a number indicates the primer begins at that intron; the number in parenthesis indicates the nucleotide at which the sequence begins; the following number indicates the sequence length; and the letter D or U indicates a downstream or upstream primer.

^bThe underlined region is the tail and the capitalized region is the sequence-specific region.

^c tT_m and cT_m represent the T_m values of the tail and the sequence-specific region of a primer, respectively.

10 ^dThe anti-sense sequences of primer #4 have 7 bp false priming sites at the 3' end.

3. Linking PCR

Linking PCR was performed as follows.

Denaturation proceeded at 94°C for 15 seconds, annealing at 55°C for 30 seconds rammed to 72°C within one minute, and then elongation at 72°C for 2-3 minutes, for a total of 15 cycles. The mixture contained 100 mM KCl, 10 mM Tris/HCl, pH 8.9, 1.5 mM MgCl₂, 50 µg/ml BSA, 0.05% (v/v) Tween 20, 200 µM of each dNTP, 1U of *Tth* (Boehringer Mannheim) and 0.1 U of Vent (New England Biolabs) DNA polymerases, 20ng each of the four DNAs, and 5 µCi of alpha-³²P-dCTP (300 Ci/mmol, Amershar) per 25 µl reaction; unless mentioned elsewhere. The PCR products were separated on a 2% agarose gel, which was then stained with ethidium bromide and UV photographed with an AlphaImager™ 2000 CCD camera (Alpha Innotech). The PCR was quantitated by PhosphorImager with ImageQuant software (Molecular Dynamics) after the dried gel was exposed for 30 minutes. The PCR yields were quantitated as "random units," i.e. the number of pixels in the PCR band minus the background.

To quantitate the accumulation of linked PQ PCR product, aliquots of the Linking PCR reaction mixture were removed from the thermocycler every 3 cycles from 9 cycles to 30 cycles. Reactions containing 40ng, 20ng, and 10ng of the four p53 DNA templates per 25 µl reaction volume (Figure 4) were used. The first appearance of the faint PQ product was dependent on the amount of DNA template in the reaction mixture, supporting the existence of four-component linking kinetics. During the later cycles, the PQ PCR product accumulated to a considerable extent, and reached a saturation point. The point at which saturation was

reached was also dependent on the amount of DNA template originally added to the reaction. Furthermore, the relative presence of intermediate products was greatly reduced after 20 cycles (Figure 4). A similar result was obtained with the F9 gene.

4. Optimization of Linking PCR parameters

a. Enzyme type, concentration, and ratio

As seen in Figure 2A, 1U of *Tth*, *Taq*, or *Tfl* was mixed with 0-0.2U *Vent* to test the yield of *p53* gene PCR product as quantitated by PhosphorImager after 15 cycles under various enzyme conditions. The results show that *Tth/Vent* in ratios of 1:0.1 and 1:0.05 (lanes 2 and 3) generated the highest yield. Relative linking PCR efficiencies were *Tth/Vent* or *Tfl/Vent* > *Tfl/Pfu* > *Taq/Pfu* > *Tth/Pfu* > *Taq/Vent*. Any single enzyme alone did not work optimally.

Further tests with *Tth* and *Vent* were performed in linking exons of the F9 gene, changing both the amounts and ratio of the two enzymes in the reaction (Figure 2B). Amounts of *Tth* ranged from 0.125U to 4U, and amounts of *Vent* ranged from 0.0125U to 0.4U, with 4U *Tth* and 0.1U *Vent* generating the highest yield (lane 7). The results show that both the absolute amount of the two enzymes and the ratio influence the efficiency of the linking polymerase chain reaction. Similar results were achieved when linking PCR was performed with segments containing 15 base pairs of complementary sequence and when 49ng of template DNA was used.

b. DNA template concentration

Using increasing amounts of each of the four DNA templates (D_1U_1 , D_2U_2 , D_3U_3 , and D_4U_4) with a constant amount of *Tth/Vent*, a "threshold" between 10 and 20ng template DNA per 25 μ l reaction volume (lanes 4 and 5) was noted. Doubling the template concentration resulted in a 14-16-fold increase in linked *p53* product. The yield of linked product is therefore the template concentration to the fourth power (yield = (total amount of all template)⁴). This was confirmed by repeating the experiment using separate steps of linking and subsequent amplification. In addition, an experiment using DNA templates with longer 15-base tails produced the same effect. Similar "threshold" effects occurred when *Taq/Pfu* and *Tfl/Vent* enzymes were applied to the F9 gene (40% GC content rather than 57% GC content as in the *p53* gene), indicating the effect is not dependent on either the enzymes or the particular DNA templates used (Figure 3B).

c. Linkage with different primer pairs

Besides the primer pairs of P/Q, other primer pairs were compared: D_1U_2 , D_2U_3 , and D_3U_4 amplified two linked regions; D_1U_3 and D_2U_2 amplified three linked regions, and D_1U_4 amplified four linked templates, respectively. The mole ratio of D_1U_2 , D_2U_3 , and D_3U_4 ; of D_1U_3 and D_2U_2 ; and of D_1U_4 (the normalized relative yield or number of potential incorporated radioactive ³²P-dCTP) reflect the relative linking efficiencies of two-, three-, and four-template reactions. Table 2 shows that Linking PCRs linking two templates are much more efficient than those linking four templates.

Also, Linking PCRs linking two templates are much less dependent on the template amount.

d. Tail length

Tails of 10, 12, or 15 bases, designed to contain
5 60-70% GC, were tested linking exons of the F8 gene.
Tails containing 12 bases were most efficient at
linkage. Further experiments with 12- and 15-base
tails in the p53 gene (T_m ranging from 21.6°C to 44.1°C)
and 12-base tails in the F9 gene (T_m ranging from 29.6°C
10 to 32.3°C) yielded the same results: 12-base tails were
most efficient.

e. Annealing temperature

The effects of annealing temperature were studied
using a Gradient Robocycler (Stratagene). For the p53
15 gene, using four templates with 12-base tails, linked
product was formed with high yields at annealing
temperatures from 50°C up to 58°C. For the F9 gene,
under the same conditions, high yields of linked
product were formed at annealing temperatures from 47°C
20 up to 55°C. The optimal annealing temperature is
relatively low and has a broad range. The optimal
annealing temperature also is associated with the GC
content of the templates.

CLAIMS

I Claim:

1. A method of linking by PCR DNA segments which occur in non-adjacent portions of target DNA wherein each DNA segment contains a sequence complementary to a sequence in the DNA segment or segments to which it is to be linked, comprising using

- a. a first primer which is complementary to the antisense strand of the first DNA segment to be linked and a second primer which is complementary to the sense strand of the last DNA segment to be linked; and
- b. at least one polymerase lacking 3'-5' exonuclease activity and at least one polymerase containing 3'-5' exonuclease activity.

2. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity is selected from the group *Tth*, *Taq*, and *Tfl*.

3. The method of claim 1, wherein the at least one polymerase containing 3'-5' exonuclease activity is selected from the group *Pfu*, *Plo*, and *Pwo*.

4. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity is present in a concentration of from about 0.125U to about 4U per 25 μ l reaction.

5. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity is present in a concentration of 4U per 25 μ l reaction.

6. The method of claim 1, wherein the at least one polymerase containing 3'-5' exonuclease activity is present in a concentration of from about 0.0125U to about 0.4U per 25 μ l reaction.

7. The method of claim 1, wherein the at least one polymerase containing 3'-5' exonuclease activity is present in a concentration of 0.1U per 25 μ l reaction.

8. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity and the at least one polymerase containing 3'-5' exonuclease activity are present in a ratio of from about 1:0.0125 to about 1:0.2.

9. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity and the at least one polymerase containing 3'-5' exonuclease activity are present in a ratio of 1:0.05.

10. The method of claim 1, wherein the at least one polymerase lacking 3'-5' exonuclease activity and the at least one polymerase containing 3'-5' exonuclease activity are present in a ratio of 1:0.1.

11. The method of claim 1, wherein the DNA segments are amplified using a first primer having a 3' portion which is complementary to the 3' end of the

5 antisense strand of the DNA segment and a 5' tail which
is complementary to the 5' end of the second primer for
the previous DNA segment; and a second primer having a
3' portion which is complementary to the 3' end of the
sense strand of the DNA segment and a 5' tail which is
complementary to the 5' end of the first primer for the
10 subsequent segment.

12. The method of claim 1, wherein the DNA
segments are amplified using a first primer having a 3'
portion which is complementary to the 3' end of the
antisense strand of the DNA segment and a 5' tail which
5 is complementary to a sequence internal to the previous
DNA segment; and a second primer having a 3' portion
which is complementary to the 3' end of the sense
strand of the DNA segment and a 5' tail which is
complementary to a sequence internal to the subsequent
10 segment.

13. The method of claim 1, wherein the DNA
segments are exons of a single gene.

14. The method of claim 1, wherein the DNA
segments are exons of different genes.

15. The method of claim 1, wherein the DNA
segments are nonexon portions of a single gene.

16. The method of claim 1, wherein the DNA
segments are nonexon portions of different genes.

17. The method of claim 1, wherein the DNA segments originate from organisms of the same species.

18. The method of claim 1, wherein the DNA segments originate from organisms of one or more different species.

19. The method of claim 1, wherein the DNA segments contain tails which are complementary to the tails of adjacent DNA segments.

20. A method of producing and amplifying DNA containing at least three linked DNA segments which occur in non-adjacent portions of target DNA, comprising

- 5 a. providing a first primer and a second primer for each DNA segment to be amplified,
 - 10 i. the first primer (termed the D primer) having a 3' portion which is complementary to the 3' end of the antisense strand of the DNA segment and a 5' tail which is complementary to the 5' end of the second primer for the previous DNA segment;
 - 15 ii. the second primer (termed the U primer) having a 3' portion which is complementary to the 3' end of the sense strand of the DNA segment and a 5' tail which is complementary to the 5' end of the first primer for the subsequent
 - 20 segment;

- b. amplifying the at least three DNA segments by multiplex PCR using the pairs of first and second primers; and
- c. subjecting the at least three amplified DNA segments to a linking PCR using a sense primer which is complementary to the antisense strand of the first segment to be linked and an antisense primer which is complementary to the sense strand of the last segment to be linked.

21. A method of producing and amplifying DNA containing at least three linked DNA segments which occur in non-adjacent portions of target DNA, comprising

- a. providing a first primer and a second primer for each DNA segment to be amplified,
 - i. the first primer (termed the D primer) having a 3' portion which is complementary to the 3' end of the antisense strand of the DNA segment and a 5' tail which is complementary to a sequence internal to the previous DNA segment;
 - ii. the second primer (termed the U primer) having a 3' portion which is complementary to the 3' end of the sense strand of the DNA segment and a 5' tail which is complementary to a sequence internal to the subsequent segment;

- 20 b. amplifying the at least three DNA segments by
 multiplex PCR using the pairs of first and
 second primers; and
- c. subjecting the at least three amplified DNA
25 segments to a linking PCR using a sense
 primer which is complementary to the
 antisense strand of the first segment to be
 linked and an antisense primer which is
 complementary to the sense strand of the last
 segment to be linked.

22. The method of claim 20, wherein after step b
and before step c unincorporated primers are removed
from the amplification reaction mixture.

23. The method of claim 21, wherein after step b
and before step c unincorporated primers are removed
from the amplification reaction mixture.

24. The method of claim 20, wherein the 5' tails
do not overlap the 3' portion which is complementary to
the 3' end of the sense strand of the DNA segment .

25. The method of claim 20, wherein the 5' tails
overlap the 3' portion which is complementary to the 3'
end of the sense strand of the DNA segment.

26. The method of claim 20, wherein the 5' tails
overlap the 3' portion which is complementary to the 3'
end of the sense strand of the DNA segment by four
nucleotides.

27. The method of claim 20, wherein the 5' tails are GC-rich.

28. The method of claim 20, wherein the 5' tails contain about 60% to about 70% G and C nucleotides.

29. The method of claim 20, wherein the 5' tails are about 10 to about 15 nucleotides long.

30. The method of claim 21, wherein the 5' tails are about 10 to about 15 nucleotides long.

31. The method of claim 20, wherein the 5' tails are 12 nucleotides long.

32. The method of claim 21, wherein the 5' tails are 12 nucleotides long.

33. The method of claim 20, wherein the DNA segments are exons.

34. The method of claim 21, wherein the DNA segments are exons.

35. The method of claim 20, wherein the DNA segments are exons of a single gene.

36. The method of claim 21, wherein the DNA segments are exons of a single gene.

37. The method of claim 20, wherein the DNA segments are exons of different genes.

38. The method of claim 21, wherein the DNA segments are exons of different genes.

39. The method of claim 20, wherein the DNA segments are nonexon portions of a single gene.

40. The method of claim 21, wherein the DNA segments are nonexon portions of a single gene.

41. The method of claim 20, wherein the DNA segments are nonexon portions of different genes.

42. The method of claim 21, wherein the DNA segments are nonexon portions of different genes.

43. The method of claim 20, wherein the DNA segments originate from organisms of the same species.

44. The method of claim 21, wherein the DNA segments originate from organisms of the same species.

45. The method of claim 20, wherein the DNA segments originate from organisms of one or more different species.

46. The method of claim 21, wherein the DNA segments originate from organisms of one or more different species.

47. The method of claim 20, wherein the linked DNA product is a copy of a gene lacking large introns.

48. The method of claim 21, wherein the linked DNA product is a copy of a gene lacking large introns.

49. The method of claim 20, wherein the linked DNA product contains a mutation.

50. The method of claim 21, wherein the linked DNA product contains a mutation.

51. The method of claim 20, wherein the linking polymerase chain reaction mixture contains a solid or liquid macromolecular additive.

52. The method of claim 21, wherein the linking polymerase chain reaction mixture contains a solid or liquid macromolecular additive.

53. The method of claim 20, wherein the macromolecular additive is polyethylene glycol.

54. The method of claim 21, wherein the macromolecular additive is polyethylene glycol.

1/9

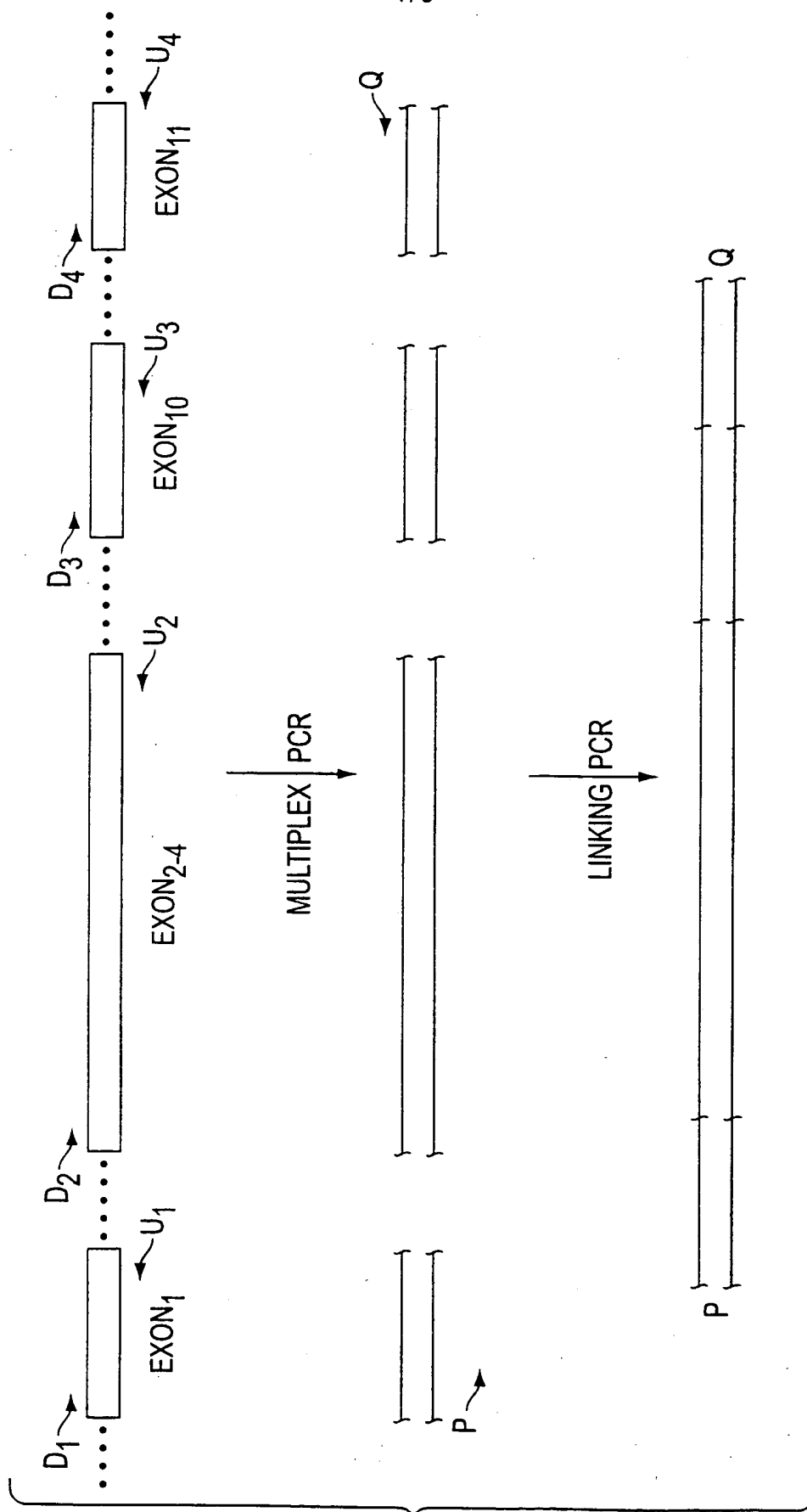


FIG. 1A

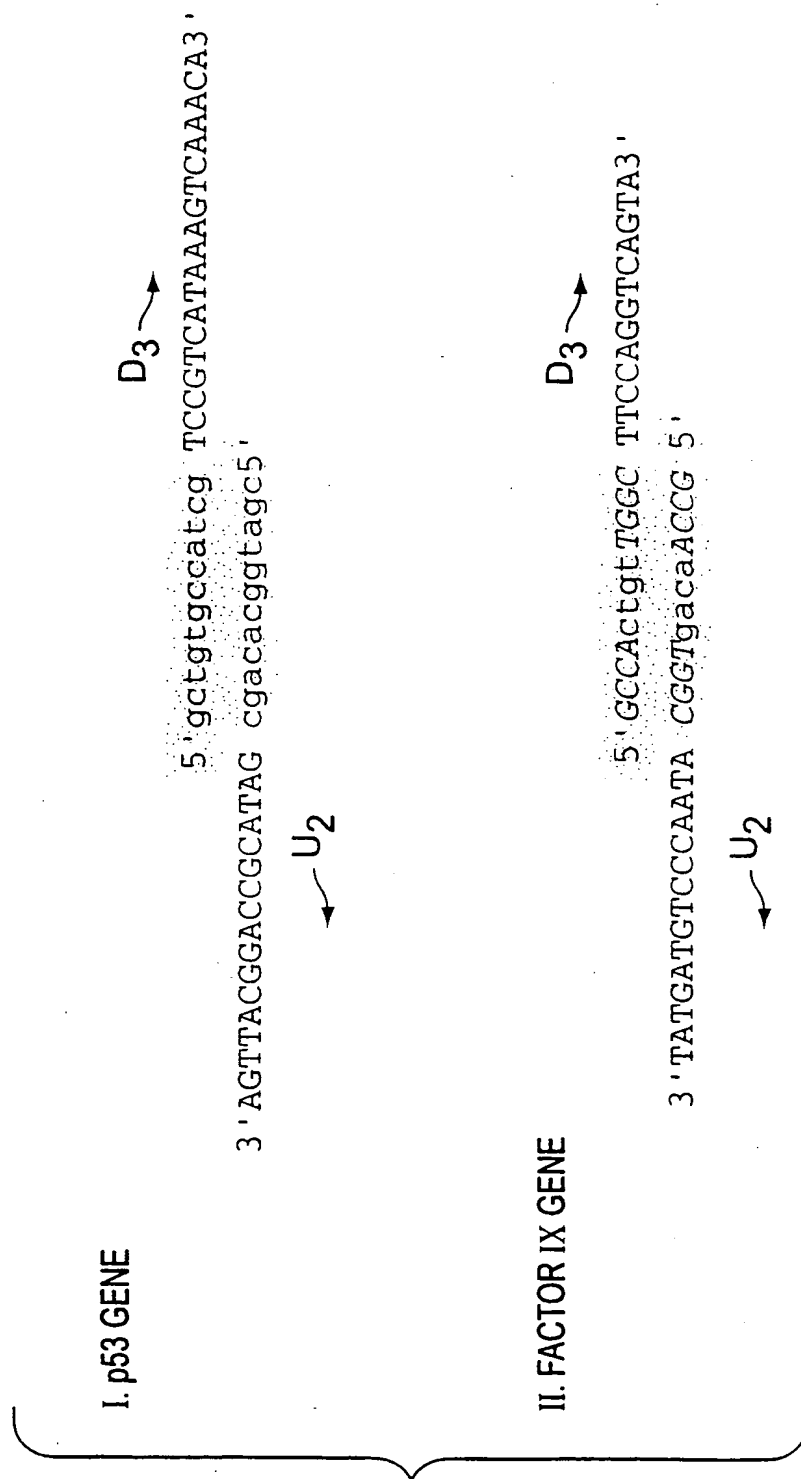


FIG. 1B

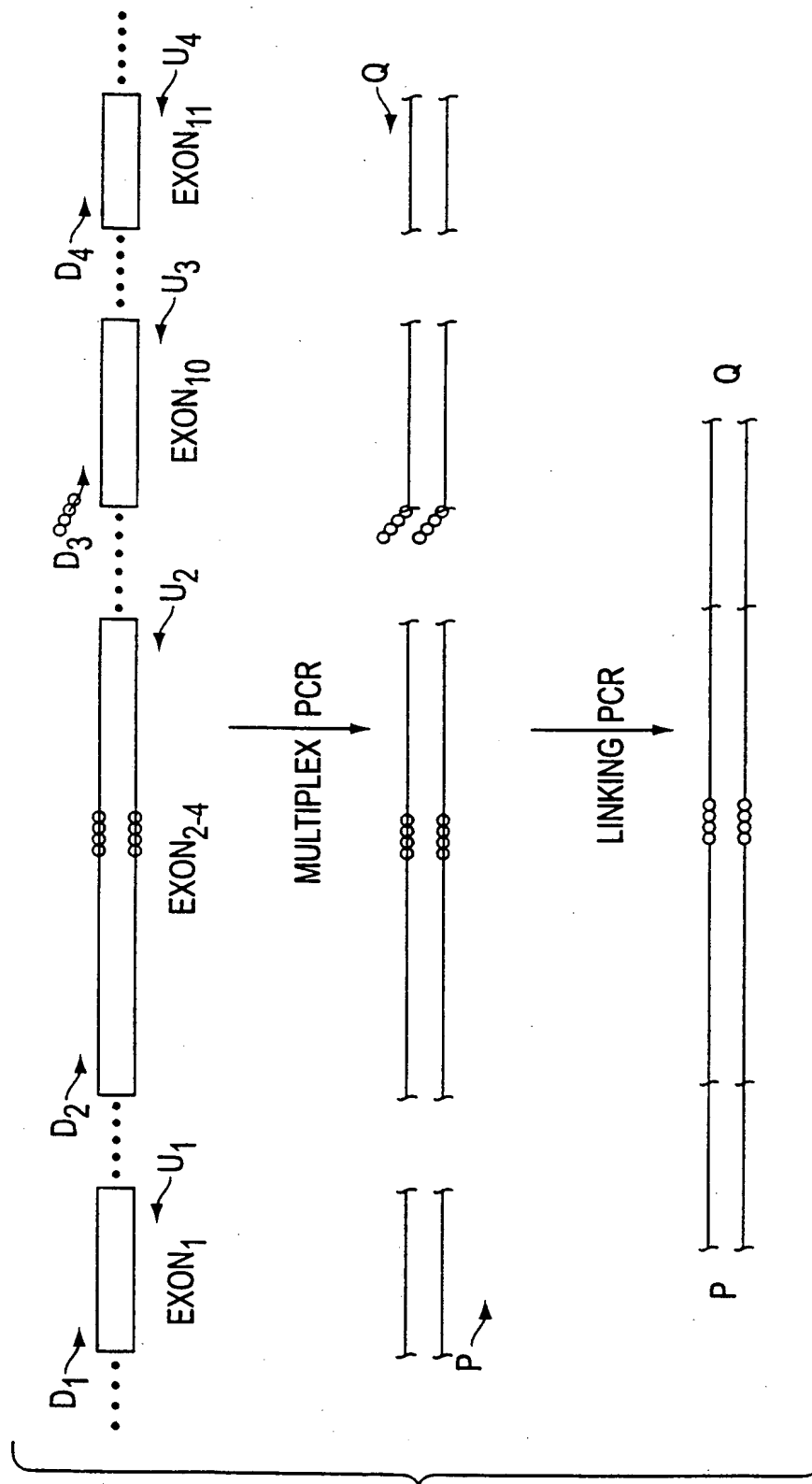
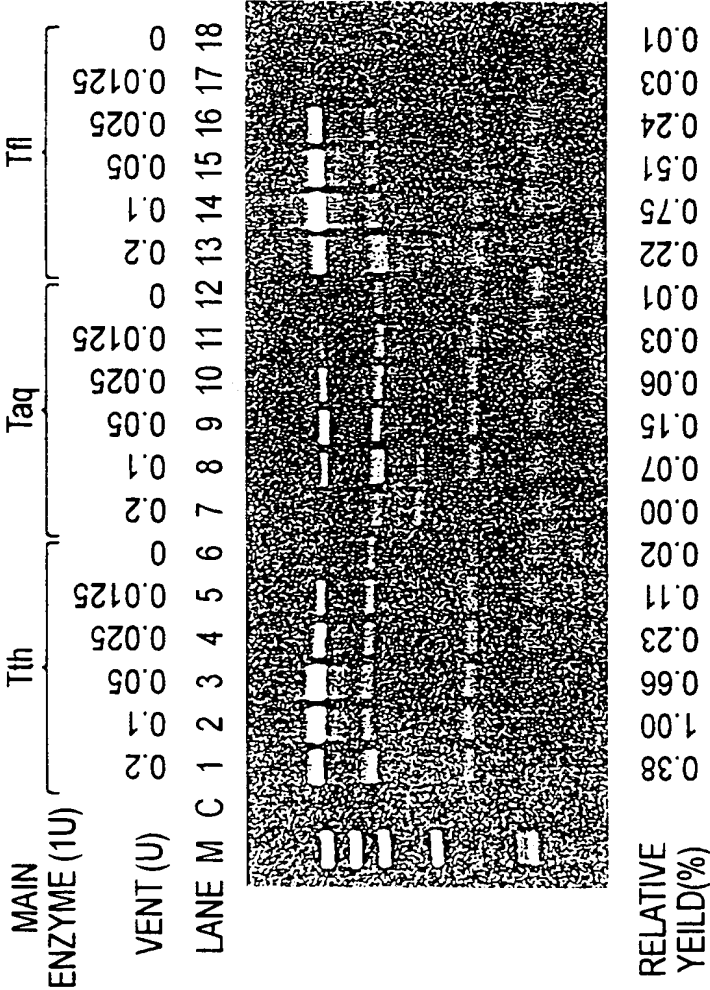
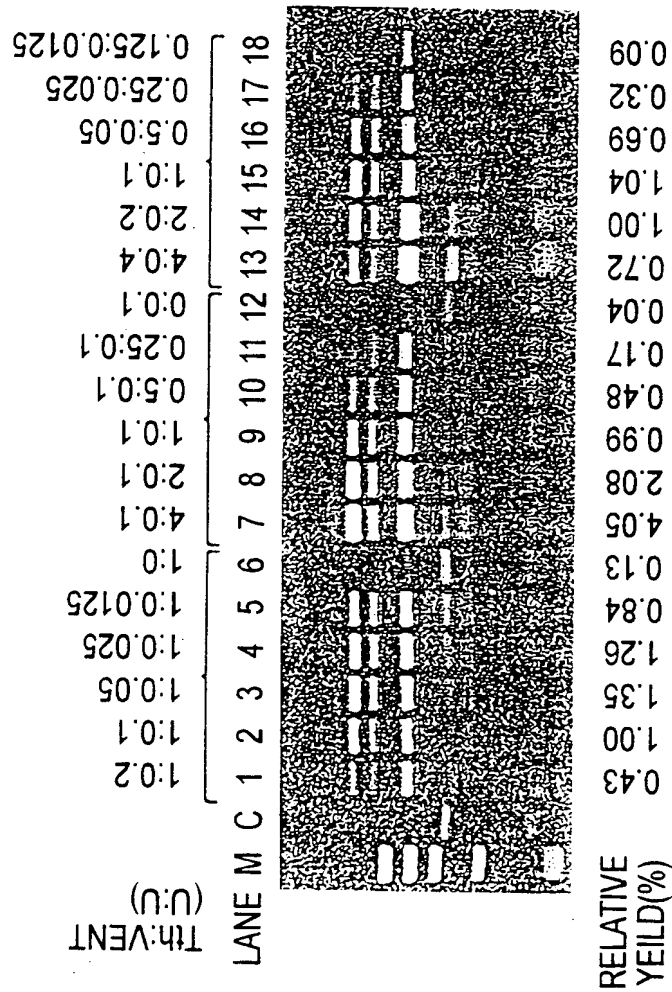


FIG. 1C





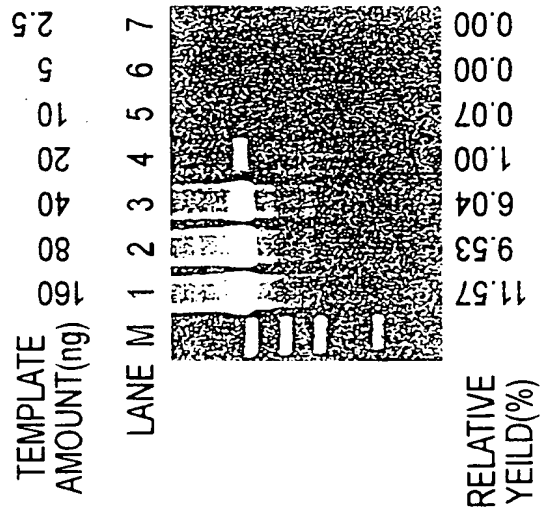


FIG. 3A

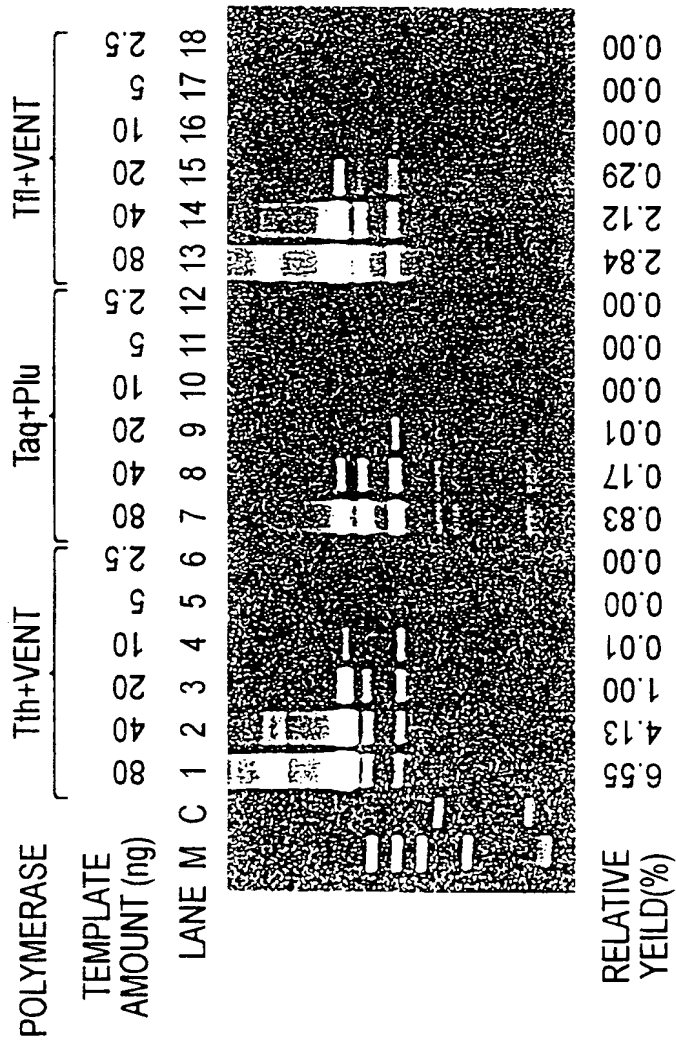


FIG. 3B

8/9

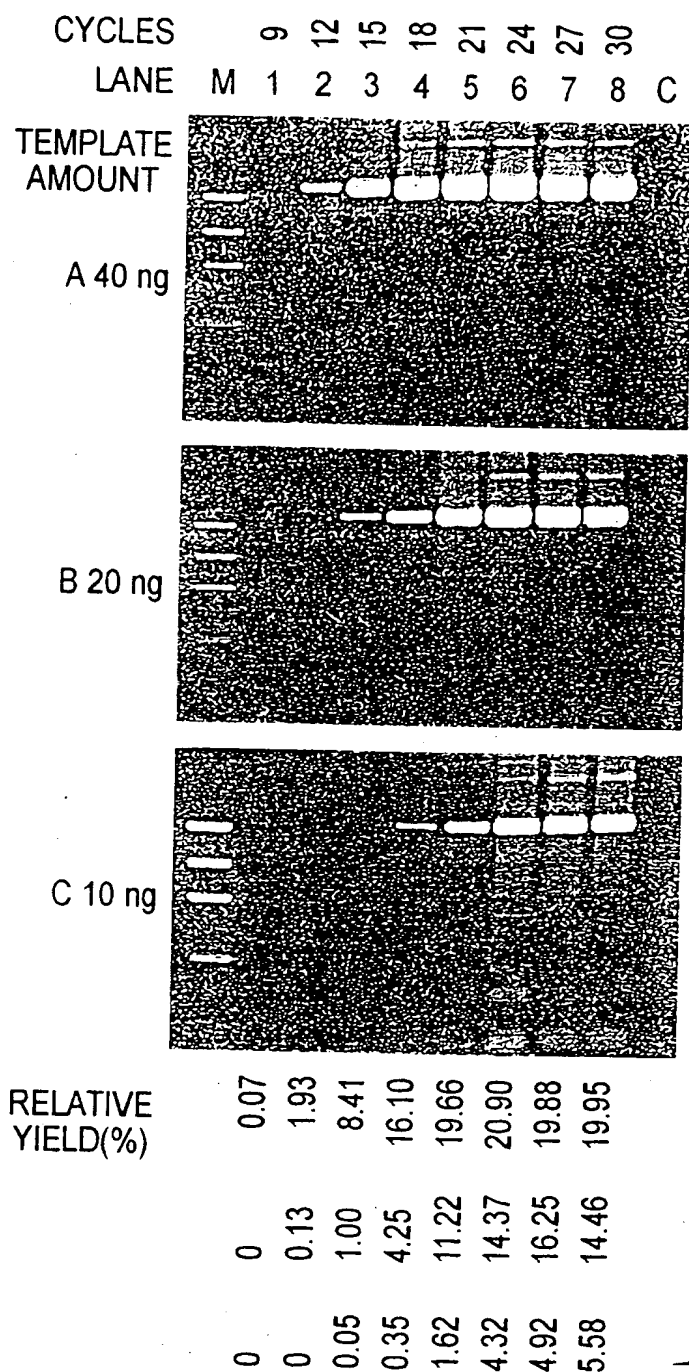


FIG. 4

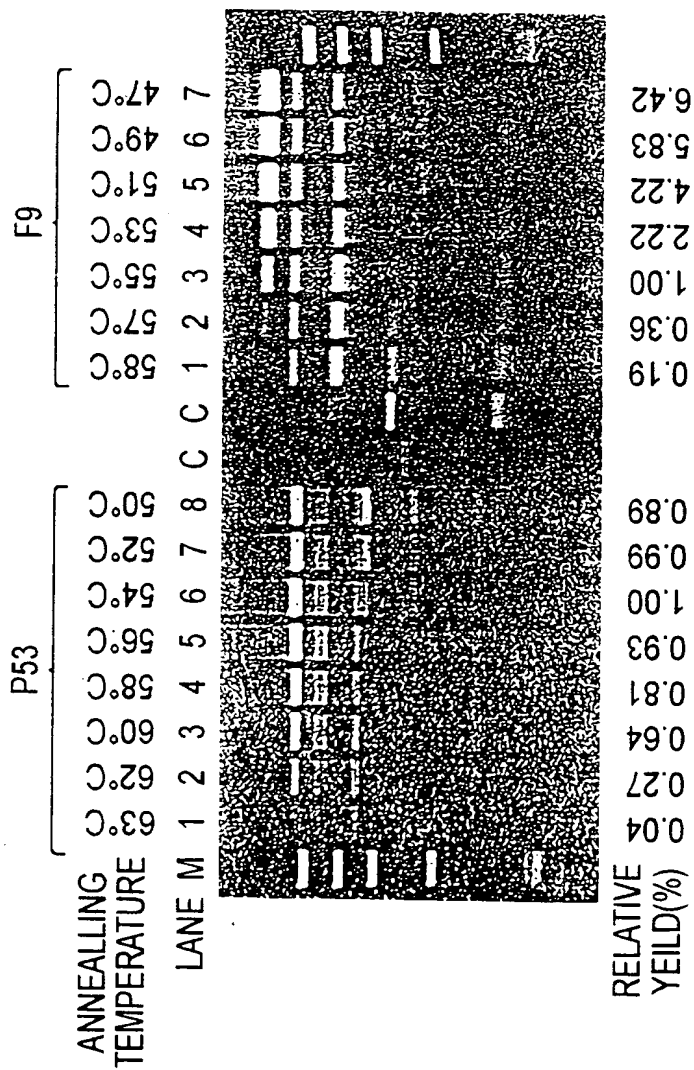


FIG. 5

SEQUENCE LISTING

<110> LIU, QIANG

<120> EXON-LINKING FOR dNA BASED DIAGNOSTICS

<130> 2124-292

<140>

<141>

<150> US 60/060319

<151> 1997-09-29

<160> 24

<170> PatentIn Ver. 2.0 - beta

<210> 1

<211> 28

<212> DNA

<213> Homo sapiens

<400> 1

gttcgcagag ggtttgtgcc aggagcct

28

<210> 2

<211> 29

<212> DNA

<213> Homo sapiens

<400> 2

aggacgaccg ctagcccgtag actcagaga

29

<210> 3

<211> 30

<212> DNA

<213> Homo sapiens

<400> 3

agcggtcgtc ctccagggtt ggaagtgtct

30

<210> 4

<211> 30

<212> DNA

<213> Homo sapiens

<400> 4

cgatggcaca gcgatacggc caggcattga

30

<210> 5

<211> 31

<212> DNA

<213> Homo sapiens

<400> 5

gctgtgccat cgtccgtcat aaagtcaaac a

31

<210> 6
<211> 31
<212> DNA
<213> Homo sapiens

<400> 6
gaggtgggtg tccctatggc tttccaacct a 31

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gacacccacc tcaccctctc actcatgtga t 31

<210> 8
<211> 29
<212> DNA
<213> Homo sapiens

<400> 8
cccgtgagga cagacccaaa acccaaat 29

<210> 9
<211> 28
<212> DNA
<213> Homo sapiens

<400> 9
caacgggtca ggaggggtg atgggatt 28

<210> 10
<211> 29
<212> DNA
<213> Homo sapiens

<400> 10
cgtgtgggtg ctgagggagg ctgtcagtg 29

<210> 11
<211> 24
<212> DNA
<213> Homo sapiens

<400> 11
tcgcagagga ggccattgga aata 24

<210> 12
<211> 25
<212> DNA
<213> Homo sapiens

<400> 12
gtaagcggtc gtgctggctg ttaga 25

<210> 13
<211> 29
<212> DNA
<213> Homo sapiens

<400> 13
cacgaccgct tactggaatt ctcttgact 29

<210> 14
<211> 27
<212> DNA
<213> Homo sapiens

<400> 14
gccaacagtg gcataaccct gtagtat 27

<210> 15
<211> 25
<212> DNA
<213> Homo sapiens

<400> 15
gccactgttg gcttccaggt cagta 25

<210> 16
<211> 28
<212> DNA
<213> Homo sapiens

<400> 16
ggtccgggat caaaggatg tttttaag 28

<210> 17
<211> 27
<212> DNA
<213> Homo sapiens

<400> 17
tgatcccgga cccatacatg agtcagt 27

<210> 18
<211> 27
<212> DNA
<213> Homo sapiens

<400> 18
acggagacag gaagcagatt caagtag 27

<210> 19
<211> 23
<212> DNA
<213> Homo sapiens

<400> 19
tcaaggagga gggagatgga cat 23

<210> 20
<211> 23
<212> DNA
<213> Homo sapiens

<400> 20
tggtgtgctt aaaatgctga agt 23

<210> 21
<211> 31
<212> DNA
<213> Homo sapiens

<400> 21
gctgtgccat cgtccgtcat aaagtcaaac a

31

<210> 22
<211> 29
<212> DNA
<213> Homo sapiens

<400> 22
cgatggcaca gcgatacgcc aggcattga

29

<210> 23
<211> 25
<212> DNA
<213> Homo sapiens

<400> 23
gccactgttg gcttccaggt cagta

25

<210> 24
<211> 27
<212> DNA
<213> Homo sapiens

<400> 24
gccaacagtg gcataaccct gtagtat

27

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/19968

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 92 07075 A (WELLCOME FOUND) 30 April 1992	20-54
Y	see whole doc, esp. claims and figures	1-19

X	US 5 023 171 A (HO STEFFAN N ET AL) 11 June 1991	20-54
	see whole doc, esp. figure 2	

Y	WO 95 16028 A (STRATAGENE INC) 15 June 1995	1-19
	see whole doc, esp. claims and example	

	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

14 January 1999

Date of mailing of the international search report

25/01/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Müller, F

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/19968

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>LIU Q. & SOMMER S. S.: "Linking PCR splices small exons into large DNA molecule amenable to rapid mutation screening" AMERICAN J. HUMAN GENETICS, vol. 61, no. 4, - October 1997 page a223 XP002089977 see the whole document -----</p>	1-54

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/19968

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9207075	A	30-04-1992	EP 0552266 A JP 6501614 T	28-07-1993 24-02-1994
US 5023171	A	11-06-1991	NONE	
WO 9516028	A	15-06-1995	US 5556772 A	17-09-1996

THIS PAGE BLANK (USPTO)